Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Journal of Theoretical Biology 526 (2021) 110761

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/yjtbi

Application of a novel mathematical model to identify intermediate hosts of SARS-CoV-2

Katherine Royce

Proof School, 973 Mission St., San Francisco, CA, United States

ARTICLE INFO

Article history:
Received 4 December 2020
Revised 17 April 2021
Accepted 7 May 2021
Available online 24 May 2021

Keywords:
Intermediate host
SARS-CoV-2
Zoonosis
Model

ABSTRACT

Intermediate host species provide a crucial link in the emergence of zoonotic infectious diseases, serving as a population where an emerging pathogen can mutate to become human-transmissible. Identifying such species is thus a key component of predicting and possibly mitigating future epidemics. Despite this importance, intermediate host species have not been investigated in much detail, and have generally only been identified by testing for the presence of pathogens in multiple candidate species. In this paper, we present a mathematical model able to identify likely intermediate host species for emerging zoonoses based on ecological data for the candidates and epidemiological data for the pathogen. Since coronaviruses frequently emerge through intermediate host species and, at the time of writing, pose an urgent pandemic threat, we apply the model to the three emerging coronaviruses of the twenty-first century, accurately predicting palm civets as intermediate hosts for SARS-CoV-1 and dromedary camels as intermediate hosts for MERS. Further, we suggest mink, pangolins, and ferrets as intermediate host species for SARS-CoV-2. With the capacity to evaluate intermediate host likelihood among different species, researchers can focus testing for possible infection sources and interventions more effectively.

1. Introduction

Zoonotic disease emergence has proven a significant threat to global health, and yet little is known about the process by which novel pathogens emerge. Zoonotic pathogens circulate in and have adapted to an animal species, the reservoir host, and are sometimes transmitted to other animal species to which they are less well-adapted, usually causing more severe illness. Such intermediate host species, generally a domestic or human-adjacent population, provide a meeting ground where evolutionary pressure and higher exposure to humans can cause a pathogen to mutate from one adapted to its reservoir host to a more human-transmissible form (Neumann et al., 2009). Coronaviruses, responsible for three spillover epidemics in humans in the 21st century, routinely emerge from their natural reservoir through intermediate host species (Corman et al., 2018), and the current pandemic provides an urgent example of the need to understand their evolution.

Bats are widely accepted as the reservoir hosts for the coronavirus family (Cui et al., 2019): coronaviruses cluster according to bat genus, and genus *Rhinopolus* carries coronaviruses known to use human ACE2, the receptor protein for both SARS-CoVs (Fan et al., 2019; Hu et al., 2017; Liu et al., 2020c). In addition, SARS-CoV-2 has been shown to replicate in bat intestinal epithelial cells (Zhou et al., 2020). However, while bats harbor an impressive diversity of coronaviruses and are a likely reservoir host for all three epidemic coronaviruses (Liu et al., 2020c), intermediate host species form a key link in the emergence of each human pathogen. Intermediate hosts are a crucial factor in the emergence of all three coronaviruses, which can adapt to different host species with ease due to their high mutation rate, large genome size, and high recombination rates during mixed infections of a single host (Li et al., 2006; Bolles et al., 2011). SARS-CoV-1 emerged into the human population through palm civets, *P. larvata* (Cui et al., 2019); although research has cast doubt on whether palm civets were a true intermediate host or merely infected by secondary transmission from initial human cases, the civet population formed a key link in the amplification of SARS-CoV-1 in humans and a resource for it to persist outside humans (Bolles et al., 2011). On the other hand, MERS required the presence of an intermediate host species, in this case dromedary camels (*C. dromedarius*), to adapt to humans, since its R_0 in humans is lower than the threshold of 1 needed to cause an epidemic (Reusken et al., 2013; Breban et al., 2013; Dudas et al., 2018). The diversity of MERS in human populations was probably due to multiple independent spillover events from camels, rather than viral mutation during human-to–human transmission, showing the importance of an

E-mail address: kroyce@proofschool.org

https://doi.org/10.1016/j.jtbi.2021.110761
0022-5193/

intermediate host species in the epidemiology of MERS (Reusken et al., 2013; Dudas et al., 2018). SARS-CoV-1 also seems to have persisted in palm civets, rather than its true reservoir host, between its spillovers in 2003 and 2004 (Shi and Hu, 2008). Intermediate host species did not show signs of obvious disease even when they tested positive for SARS-CoV-1 (Guan et al., 2003), supporting the theory that mutation in intermediate hosts is driven by adaptation to that species and thus that such hosts play a crucial role in emergence of human-transmissible pathogens. Active animal markets have been linked to the emergence of all three epidemic coronaviruses, showing the importance of intermediate host species in the amplification, recombination, and transmission of each pathogen (Liu et al., 2020c). It is therefore important to identify possible intermediate host species quickly and accurately to support efforts in controlling outbreaks.

While currently the forefront of public health efforts, coronaviruses are not the only zoonoses which emerge through intermediate hosts; indeed, this mode of emergence may be the norm for animal-derived diseases (Cunningham et al., 2017). In this paper, we present a model which ranks potential intermediate host species for any given pathogen, given genus-level information about the reservoir host and ecological observations for potential intermediate host species, and use epidemiological data in humans to validate the model in the case of SARS-CoV-1 and MERS. Few mathematical models capture the behavior of emerging infectious diseases through the entire course of their evolution, despite the demonstrated importance of this method of pathogen emergence and the predictive capacity of such models (Lloyd-Smith et al., 2015; Allen et al., 2012). Building on a preexisting mathematical model for the emergence of infectious zoonoses through an intermediate host (Royce and Fu, 2020), we have linked the spread of disease from the reservoir species, through the intermediate host population, and into humans to factors specific to a particular intermediate species, allowing us to compare the severity of outbreaks that result from pathogen evolution through different intermediate host species. We classify intermediate host species based on six parameters: their biological similarity to, contact with, and transmission risk among both humans and the reservoir host. The major innovation introduced here is the ability to predict intermediate host species, by making the interspecies transmission parameters functions of intermediate-host-specific qualities, such as pathogen-receptor compatibility and observed rates of contact with humans and reservoir hosts.

Applying the model to the outbreaks of epidemic coronaviruses, we find that it accurately identifies palm civets as the intermediate host of SARS-CoV-1 in 2003 and dromedary camels as the intermediate host of MERS-CoV in 2012. While researchers have suggested multiple candidates for the intermediate host of SARS-CoV-2, these attempts are inconclusive and the subject of current debate (Xu et al., 2020; Yuan et al., 2020; Liu et al., 2020c); we therefore apply the model to the current pandemic to identify mink, pangolins, and ferrets as the most likely intermediate hosts of SARS-CoV-2. While we have focused on coronaviruses, this model can be used to predict intermediate host species for any zoonotic pathogen given data on the reservoir species and the interactions between potential intermediate hosts and other populations. As such, it provides a more rigorous and less costly way to identify intermediate host species than sampling animal populations, and can focus research into emergence processes more efficiently. Emerging and endemic zoonoses are, in general, under-studied, and many such pathogens could be explored using the model presented here to provide valuable understanding where lab and field resourdes are limited. This findings can help prioritize research into the origins of the current pandemic, and provide a mathematical framework for investigation of intermediate host species of many important pathogens.

2. Methods

We have modified the introductory intermediate host model presented by Royce and Fu (2020) to better model epidemic spread based on factors intrinsic to the intermediate host species. The full list of ten differential equations, grouped into three linked SIR models, is reproduced for convenience in Tables 1,2, and a full justification of the system can be found in the original paper (Royce and Fu, 2020). This earlier analysis computes the global R_0 of the pathogen as the maximum of each species R_0 using a next-generation matrix, and introduces a quantitative tool for analyzing transmission dynamics that include intermediate host species.

Firstly, where the original model did not distinguish between disease dynamics of the two strains, the wild and human-transmissible strains of the pathogens now have different transmission rates (β_{iw} and β_{im}) in the intermediate host species, better representing a pathogen evolving to fit a new host species over the course of an outbreak. Further, our modified framework specifies the transmission parameters involving pathogen evolution as it progresses through species-transmission rates β_{iw} and β_{im} for both strains of the pathogen in the intermediate host species and humans, the spillover rates t_i and t_h between species, and the mutation rate μ of the pathogen to a human-transmissible form-as functions of intermediate host species-specific parameters. Together, these changes make the model responsive to qualities unique to a given intermediate host species: where the original model did not have the capacity to include intermediate host species characteristics, this revision links global disease dynamics to the pathogen's progress through the intermediate host, enabling the comparison of different potential host species.

We define the intermediate host species parameters (Table 3) based on the biological, ecological, and behavioral factors governing possible pathogen transmission between the candidate intermediate host and the reservoir species or humans, resulting in six parameters specific to the intermediate host species. These parameters investigate pathogen transmission from a reservoir species, here assumed to be *Rhinopolus* bats, or to humans, and thus only consider factors pertinent to the particular context of intermediate host interactions with that species. For example, to calculate the risk of transmission to humans, ρ_h, for a species such as palm civets in the context of SARS, we consider civets to be primarily farmed for fur and thus a typical human-civet interaction to involve a high risk of blood exchange, even though the typical human may never interact with a civet other than through binoculars.

When applying our methods to the coronaviruses under consideration, we assume horseshoe bats (*R. sinicus*) are the reservoir host, since this species is known to harbor SARS-CoV-1 (Xu et al.,

Table 1
The ODE systems of the intermediate host model, edited to include two distinct pathogen strains spreading among intermediate hosts. The model simulates a pathogen spreading between susceptible (S), infected (I), and recovered (R) individuals in a reservoir, intermediate, and human population, indicated by the subscripts *r*, *i*, or *h*. We simulate the pathogen mutating to a human-transmissible form, resulting in a class T_i of infected intermediate hosts who can transmit the disease to humans.

Reservoir	$dS_r/dt = b_r - \beta_r S_r I_r - m_r S_r$ $dI_r/dt = \beta_r S_r I_r - \gamma_r I_r - m_r I_r$ $dR_r/dt = \gamma_r I_r - m_r R_r$
Intermediate	$dS_i/dt = b_i - \beta_{iw} S_i I_i - t_i S_i I_r - \beta_{im} S_i T_i - m_i S_i$ $dI_i/dt = \beta_{iw} S_i I_i + t_i S_i I_r - \mu I_i - \gamma_{iw} I_i - m_i I_i$ $dT_i/dt = \mu I_i + \beta_{im} S_i T_i - \gamma_{im} T_i - m_i T_i$ $dR_i/dt = \gamma_{iw} I_i + \gamma_{im} T_i - m_i R_i$
Humans	$dS_h/dt = b_h - \beta_h S_h I_h - t_h S_h T_i - m_h S_h$ $dI_h/dt = \beta_h S_h I_h + t_h S_h T_i - \gamma_h I_h - m_h I_h$ $dR_h/dt = \gamma_h I_h - m_h R_h$

Table 2
Parameter definitions.

S_x	susceptible individuals of species x
I_x	infected individuals of species x
T_i	intermediate hosts infected with human-transmissible strain
R_x	recovered individuals of species x
β_r	transmission rate among reservoir host
β_{iw}	transmission rate of wild strain among intermediate host
β_{im}	transmission rate of human-transmissible strain among intermediate host
β_h	transmission rate among humans
γ_x	recovery rate among species x
b_x	birth rate among species x
m_x	natural mortality rate among species x
t_i	transmission rate from reservoir to intermediate hosts
t_h	transmission rate from intermediate hosts to humans
μ	mutation rate of the pathogen in the intermediate host population

Table 3
The intermediate host (IH) species parameters. Each of the six parameters is calculated from a simple similarity matrix for each potential IH species (see Table 5). σ_i is calculated by considering biological factors affecting the susceptibility of the species to the pathogen, κ_i by considering the ecology of the species in question, and ρ_i by evaluating the typical role the species plays in its natural environment or its use by humans.

Parameter	Definition	Meaning	Units
σ_r	reservoir similarity	biological similarity of the IH to the reservoir species	unitless
σ_h	human similarity	biological similarity of the IH to humans	unitless
κ_r	reservoir contact	contact between IH individual and reservoir species per day	contacts/day
κ_h	human contact	contact between IH individual and humans per day	contacts/day
ρ_r	reservoir risk	risk of pathogen transmission per contact between IH and infected reservoir host	infections/contact
ρ_h	human risk	risk of pathogen transmission per contact between human and infected IH	infections/contact

2020; Lau et al., 2005), may harbor SARS-CoV-2 (Xu et al., 2020; Li et al., 2019), and is at least similar in key biological and behavioral aspects to the bat reservoir host of MERS, the Egyptian tomb bat *T. perforatus* (Yuan et al., 2020). The original intermediate host model assumed that the transmission and recovery rates (and thus the reproduction number) of the pathogen in both animal species was the same. Here, we have revised that assumption to reflect a pathogen evolving based on the host species in which it finds itself: the transmission rate of the wild strain of the pathogen, the original spillover from the reservoir host, is scaled by a factor of σ_r, simulating a pathogen that spreads less easily in a new species. We assume that once the pathogen evolves to match its new host, it regains its former transmission rate ($\beta_{im} = \beta_r$). We have not modified the recovery rates in the intermediate host species, γ_{iw} and γ_{im}, since a similar scaling would cancel in the calculation of $R_0 = \frac{\beta}{\gamma}$ and produce an effectively unchanged pathogen strain. While we initially modeled the transmission rate of the pathogen in humans by scaling β_{im} by a factor of σ_h, this choice led to SARS and MERS epidemics with values of R_0 higher (4.86 and 1.04) than those seen in reality. Setting $\beta_h = \sigma_h\beta_{iw} = \sigma_h\sigma_r\beta_r$ produced epidemics with smaller and more realistic values of R_0, and so we used that formulation. Mathematically, this change occurs because R_0, as shown in Eq. (1) below, can be written in terms of β_r only, and since $\beta_h b_h = \sigma_r\sigma_h\beta_r b_h < \sigma_r\beta_r b_h$, this choice for β_h produces lower values for R_0. Biologically, although this framework gives the transmission rate of the pathogen in humans in terms of the transmission rate of the wild strain rather than the mutant, the added factor may represent the challenges inherent to adapting to multiple new host species in a timespan of months.

We assume the transmission rate between species is directly proportional to both the contact and risk of transmission between species, and thus define the transmission rates t_i and t_h as the product of the contact and risk parameters for the appropriate species ($t_i = \kappa_r\rho_r$ and $t_h = \kappa_h\rho_h$, respectively). Finally, we define the mutation rate of the pathogen to a human-transmissible form while circulating in an intermediate host species as directly proportional to the animal host's similarity to, contact with, and risk of transmission to humans ($\mu = \sigma_h\kappa_h\rho_h$), since a pathogen already adapted to a human-like host is more likely to become transmissible among humans, and intermediate host-human transmission creates a selective pressure for this mutation to occur. Table 4 summarizes the definitions for each modified transmission parameter.

To calculate the intermediate host species parameters, we assign a binary score for each of the questions summarized in Table 5. The questions used in the calculation of similarity involve the role the intermediate host plays in its ecosystem, as a pathogen is likelier to succeed in organisms that fill similar niches as its original host, as well as whether the given strain of the pathogen (wild or human-transmissible) can effectively enter and replicate in intermediate host cells. Those used in the calculation of contact measure ecological factors, such as whether the intermediate host shares a natural habitat with bats or humans, whether it is kept domestically, whether it shares dietary habits with bats or humans, and whether it shares its time of activity with the population of interest. Finally, the similarity scores for risk consider if a typical contact between an individual intermediate host and a bat or human involves any of the typical vectors of disease transmission: blood; excrement; bites; saliva, a factor in the spread of Nipah virus from bats to pigs (Cunningham et al., 2017); enclosed spaces, in which an airborne virus can more easily spread; or, in the case of humans, milk (a factor in the spread of MERS from camels to humans (Omrani et al., 2015)). For example, if the potential intermediate host species is farmed for fur or food, we assume most contact with humans involves blood and occurs in an enclosed space, whereas if it is a wild species, we assume contact with humans is limited to occasional contamination. Each parameter is calculated from the n questions by summing 0.01 as a baseline level of similarity and $\frac{1}{n} - 0.01$ to the total. For example, for σ_r, which is scored from 3 questions, we assume the candidate IH species is at least 1% similar to the reservoir species (since they share a kingdom), and then add 0.33 to the total for each positive answer. This method assumes that each factor is equally important to the success of the pathogen and ensures that each parameter takes a value in the half-open interval (0,1].

To investigate the intermediate hosts of the three coronavirus epidemics, we searched the literature for investigations into animal hosts for each pathogen. Potential intermediate hosts for SARS-CoV-1 were found in Li et al. (2006) and Shi and Hu (2008), while those for MERS were found in Reusken et al. (2013) and

Table 4
The transmission parameters defined as functions of intermediate host species parameters. Note that each transmission parameter has units of infections/day, although the type of individual and infection varies.

Parameter	Function
β_{iw}	$\sigma_r\beta_r$
β_{im}	β_r
β_h	$\sigma_h\beta_{iw}$
t_i	$\kappa_r\rho_r$
t_h	$\kappa_h\rho_h$
μ	$\sigma_h\kappa_h\rho_h$

Table 5
The questions scored for intermediate host (IH) species parameters. The questions are scored both for reservoir-IH interactions and for IH-human ones.

Similarity	Contact	Risk
does the IH share a niche with bats/humans?	does IH share its natural habitat with bats/humans?	does contact include blood?
can bat/human pathogen use IH receptor?	is the IH wild/domestic?	does contact include excrement?
can bat/human pathogen replicate in IH cells?	does the IH share dietary habits with bats/humans?	does contact include biting?
	is the IH active at the same time as bats/humans?	does contact include saliva?
		does contact take place in enclosed spaces?
		does contact include milk? (ρ_h only)

Omrani et al. (2015). Although an area of active research at the time of writing, potential intermediate hosts for SARS-CoV-2 were found in Yuan et al. (2020), Liu et al. (2020a,b,c), Shi et al. (2020), Ji et al. (2020) and Lam et al. (2020). The similarity matrices for each simulation were calculated based on the context of the epidemic-for example, after the 2003 SARS epidemic, palm civets are much less likely to be farmed, so their contact with humans is less likely in 2020 than in 2002 (Shi and Hu, 2008)-and the place of emergence, resulting in different intermediate host parameters, even for the same species, for different epidemics (see the Supplementary Information for the data for each coronavirus and the code used to run the simulations). For SARS-CoV-1, we assume $\beta_r = 0.25$ and $\gamma_r = 0.14$, reflecting a moderately transmissible disease ($R_0^w = 1.78$) in its reservoir host (Chowell et al., 2004). For MERS, we could not find an explicit source for β_r, but since bats transmit betacoronaviruses to individuals of other bat species at a rate of 0.002 infections per unit time (Latinne et al., 2020), we assume bats are 40 times more likely to transmit to members of their own species and thus set $\beta_r = 0.08$. We leave γ_r unchanged, resulting in an $R_0^w = 0.57$. We then ran the model with a fixed initial proportion of 10% infected reservoir hosts, matching the equilibrium proportion of infected wild animals, and simulated the spread of the pathogen through each intermediate host population. We implemented a parameter sensitivity analysis for the reservoir species input parameters (see Table 7), varying a parameter x in the interval $[0.01, 3x]$ using a step of $\frac{x}{10}$. The baseline values x for each parameter are given in the text (for parameters introduced here) or in Royce and Fu (2020), and the range for each parameter in Table 7.

The intermediate host species were ranked using an unweighted average of the maximum proportion of infected humans and the pathogen's global R_0. We considered the maximum, rather than equilibrium, proportion of infected humans because a pathogen that peaks at a higher proportion of infected humans in its initial spillover population (such as the vendors at an animal market) has a higher chance of spreading among the human population generally even if it eventually reaches a lower equilibrium, and because in all our simulations a higher maximum proportion of infected humans was correlated with a higher equilibrium proportion. The global R_0, calculated for each potential intermediate host species using a next-generation matrix, is the maximum of the basic reproduction of each strain in each species,

$$R_0 = \max\left\{\frac{\beta_r b_r}{m_r(\gamma_r + m_r)}, \frac{\beta_{iw} b_i}{m_i(\gamma_{iw} + m_i)}, \frac{\beta_{im} b_i}{m_i(\mu + \gamma_{im} + m_i)}, \frac{\beta_h b_h}{m_h(\gamma_h + m_h)}\right\}, \quad (1)$$

and measures the epidemic spread in humans and animal populations, where the pathogen is assumed to mutate (see Royce and Fu, 2020; Van den Driessche and Watmough, 2002). In this framework, R_0 directly depends only on the similarity parameters, as these control the transmission rates. However, other intermediate host parameters control the speed with which the epidemic spreads in the intermediate compartment and the rate of mutation, as well as the initial level of exposure humans face, and thus also affect the epidemic in humans. The global R_0 and maximum proportion of infected humans, in addition to being the traditional measures of epidemic severity, are positively correlated with other measures of interest, such as the time to initial spillover to humans, the time to the epidemic peak in humans, and the equilibrium proportion of infected humans, although they do not always correspond to the most severe outbreak in the intermediate host species.

3. Results

We accurately identified palm civets as the most likely intermediate host of SARS-CoV-1 and dromedary camels as the most likely intermediate host of MERS. For SARS-CoV-1, passage through palm civets produced a pathogen that infected a maximum of 58% of the human population in 49 days, with an R_0 of 2.97. This simulation matches most estimates for the R_0 of SARS-CoV-1 without control measures, and the epidemic produced by our simulation matches the epidemic progress in 2003 (Chowell et al., 2004; Banos and Lacasa, 2007). For MERS, passage through dromedary camels produced a pathogen that infected a maximum of 26% of the susceptible human population in 56 days from its emergence from the reservoir host, with an R_0 of 0.64. This simulation matches the epidemic progress of MERS, which had a human $R_0 = 0.69$ and took many months to emerge in several independent spillover events (Reusken et al., 2013; Dudas et al., 2018).

We identified mink (*M. lutreola*), pangolins (*M. pentadactyla*), and ferrets (*M. furo*), in descending order, as the most likely intermediate hosts of SARS-CoV-2. Passage through mink produced a pathogen that infected 56% of humans 48 days after initial emergence, while passage through pangolins produced a pathogen that infected 50% of humans 44 days after emergence. Both simulations had an R_0 of 3.17 in humans, which matches stated ranges for the reproduction number of SARS-CoV-2 (Liu et al., 2020c; Zhang et al., 2020). The success of mink as an intermediate host resulted from the high levels of risk associated with fur farming, while that of pangolins and ferrets from their noted biological similarity to humans and bats in the context of infection with SARS-CoV-2 (both wild and transmissible strains of the virus can replicate and transmit in their cells). All candidate intermediate host species that produced an epidemic that infected more than 45% of susceptible humans or had an $R_0 > 2$ are summarized in Table 6. Interestingly, they are all small carnivores or snakes.

For each of the three epidemic simulations, there were several clusters of species causing epidemics of similar severity (for example, passage through raccoon dogs, ferrets, pangolins, and mink caused a similar number of infected humans in an almost identical timespan in the SARS-CoV-2 simulations), so our results may not definitively identify one particular intermediate host species absent better data. The rankings our model produces are robust to a parameter sensitivity analysis for all reservoir species parameter inputs $\beta_r, \gamma_r, b_r, m_r$ (Table 7). (Values of β_r between 0.01 and 0.085 gave an identification of racoon dogs as the most likely intermediate host for SARS-CoV-2; however, these values give an R_0 in the reservoir species of 0.07 to 0.61 and are thus unlikely to match

Table 6
Potential intermediate hosts for SARS-CoV-2 ranked from highest to lowest by severity of epidemic.

Species	Final I_h	Days to Spillover	Maximum I_h	Days to Maximum	R_0	Rank
mink	0.225	23	0.559	48	3.17	29.55
pangolin	0.211	22	0.499	44	3.17	26.52
ferret	0.217	22	0.508	46	2.12	26.48
raccoon dog	0.210	23	0.480	48	2.12	25.06
civet	0.199	24	0.447	45	3.17	23.91
Chinese cobra	0.199	24	0.444	44	3.17	23.79
many-banded krait	0.191	25	0.415	45	3.17	22.34
cat	0.211	26	0.455	58	1.61	23.57
coyote	0.173	24	0.329	51	2.12	17.52

Table 7
Results of a parameter sensitivity analysis for the reservoir species input parameters, showing the most likely intermediate host species and the parameter range that produced that result (if different than those identified above) or a ✔ if the analysis agreed with our results for all values of the parameter. The possible ranges for b_r and m_r were [0.010, 0.185]; for β_r, [0.010, 0.735] (SARS and COVID) and [0.010, 0.234] (MERS); and for γ_r, [0.010,0.416].

Disease	b_r	m_r	β_r	γ_r
SARS	pig [0.119, 0.185]	pig [0.010, 0.035]	pig [0.385, 0.510], cat [0.535,0.735]	pig [0.010, 0.094]
MERS	✔	✔	✔	✔
COVID	✔	✔	raccoon dog [0.010, 0.085]	✔

real transmission rates.) These results suggest that while the discrepancies between the epidemics produced through passage in different species may be small in their effect in humans, they are still meaningful.

4. Discussion

This model correctly identifies the intermediate host species of SARS-CoV-1 and MERS, showing that it is possible to predict the intermediate host species of a given pathogen using only a genus-level identification of the reservoir host, data regarding pathogen spread in that population, and species data for potential intermediate hosts. Making the disease transmission parameters functions of species parameters in a traditional SIR model has been theorized to allow more accurate modeling of disease dynamics, and our research supports this theory (Lloyd-Smith et al., 2015; Allen et al., 2012).

The intermediate host species parameters presented here are general indicators of similarity, contact, and risk of pathogen transmission between the given intermediate host species and humans or bats. While a more detailed framework for quantifying the risk of spillover through each intermediate host species certainly could be devised, the ease of computing each parameter in this framework, together with the accuracy of its results, suggest that this model usefully captures the essential details of spillover.

Our results match the two previous coronavirus epidemics in good detail. Our simulation of SARS-CoV-1 in palm civets matches the epidemic's R_0 and case burden in humans. In the case of MERS, our model assumes sustained transmission once the human-transmissible strain of the pathogen enters the human population, and thus our simulation achieves a nonzero endemic equilibrium in humans while MERS did not transmit effectively outside of hospitals and family clusters (Dudas et al., 2018). However, our model focuses on the initial cases of an emerging zoonosis, as identifying potential public health threats before they become outbreaks or epidemics is more useful in the long term than judging severity based only on outcomes (Royce and Fu, 2020). Thus, we assign comparatively less importance to the progress of the epidemic after the pathogen establishes a foothold in humans. Although our simulations of SARS-CoV-1 and MERS are reasonably accurate, they may be improved by adding more species-specific reservoir host data, as our model assumed *Rhinopolus* bats as the reservoir host for all three cases. Further, since our reference for the β_r of MERS in its reservoir population is based on transmission rates between individuals of different bat species, we suggest a more detailed inquiry into transmission rates of specific pathogens in specific reservoir species, rather than among wild bats as a uniform population. Such analysis could clarify if R_0^w for MERS is indeed below the traditional threshold for epidemic success.

Our results for MERS in particular may be subject to the critique that preexisting knowledge of the intermediate host species' identity may have influenced our setup, and it is true that knowledge of the hypothesized role that camel milk played in the transmission of MERS led us to include that question in Table 5. However, milk consumption is a valid possible factor for disease transmission to humans from a variety of animals, including camels, cows, donkeys, and goats, and we included these species in our analyses for all three coronaviruses. We also included the question about milk consumption in our construction of the similarity matrices for all three epidemics. Had the consumption of milk been a determining factor for the identification of camels as intermediate hosts, we would expect the model to rank other milk-producing species as higher probable intermediate hosts for the other coronavirus epidemics. Instead, passage through cattle, which had milk production as a risk factor for all three diseases, produced a Covid-19 strain with an R_0 of 1.26, a MERS strain with an R_0 of 0.40, and a SARS strain with an R_0 of 1.26, suggesting that including this parameter did not unduly weight our analysis.

Our research suggests ferrets, pangolins, and mink as potential intermediate hosts for SARS-CoV-2; more broadly, it identifies small mammalian carnivores and carnivorous snakes as species of interest in its spread. Since in humans and other carnivores, coronaviruses mainly seem to cause respiratory illnesses, while in livestock they seem to cause gastrointestinal illnesses (Corman et al., 2018), this finding matches the observed cross-species presentation of disease. While this theoretical finding may seem to conflict with early reports that pangolins served as an intermediate host species-passage through mink produced an epidemic that infected 5% more of the human population in a comparable amount of time-our data support the initial findings of Liu et al. (2020a). Pangolins lack the ecological similarity to or close contact with humans that made civets (small carnivores farmed for fur) and dromedary camels (domestic animals kept for racing and milk) effective amplifying hosts for SARS-CoV-1 and MERS, and it may be only an effect of the scrutiny the species faced early in the outbreak that there is data for pangolin receptivity to SARS-CoV-2 that we

had to infer for species less tied to the pandemic. If we assume that SARS-CoV-2 required passage through an intermediate host to achieve successful adaptation to humans, as both previous coronavirus epidemics did, this model suggests probable candidates for the role of amplifying species independent of early-reporting bias.

The questions that serve as the basis of the similarity matrices for each intermediate species parameter, as well as the values assigned to them for each candidate species, can surely be improved with more detailed surveys of potential intermediate host species, and we strongly recommend more detailed research in this area. Further, it is possible that different researchers could assign different answers to the binary questions that form the foundation of our similarity matrices; while this possibility may introduce some subjectivity into the scoring metrics, we feel the simplicity of use compensates for the possibility of small changes in scoring. Mathematically, this model raises the question of why a correlation exists between epidemic severity measures such as the global R_0 and maximum proportion of humans infected, which are generally regarded as the mathematical determinants of the intensity of an epidemic, and the time to spillover and maximal infection and the endemic proportion of humans infected. While a connection between the two may seem obvious from an epidemiological point of view, there is not an immediate mathematical reason that they should be linked. Further, we would prefer to give the transmission rate in humans, β_h, in terms of β_{im}, and we plan to research a biological explanation of why the choice $\beta_h = \sigma_h \beta_{iw}$ matches previous epidemics more exactly. Intriguingly, in our simulations (see the Supplementary Information), successful epidemics in humans seem to fade in the intermediate host species, which may be a reflection of a successful adaptation to a new host. We have not accounted for recovered individuals in the reservoir host population at the beginning of our simulations, reflecting a lack of data into prevalence of bat coronaviruses as well as indications that viral outbreaks in the reservoir species may cycle seasonally with births of newly susceptible individuals (Dudas et al., 2018). Finally, we plan to test this model with other zoonoses that emerged through an intermediate host, such as Nipah and avian influenza. With better data, we hope to replicate our success at identifying intermediate hosts for other emerging infectious diseases.

5. Conclusion

We have developed a mathematical model that can rank the likelihood of intermediate host species based solely on species-level contact parameters for each candidate intermediate host, and genus-level identification of a reservoir host, using human epidemic data as a point of comparison. Here, we show that the model accurately identifies the intermediate host species of SARS-CoV-1 and MERS, and use it to identify potential intermediate host species for SARS-CoV-2. Our results confirm the focus on small carnivore species as important hosts in the emergence of pathogens-and coronaviruses in particular-and offer guidelines for research into the emergence and control of the current pandemic.

By identifying species through which pathogens emerge relatively easily, as well as those which have some biological or behavioral resistance to emerging infections, epidemiologists can identify nonmedical interventions helpful in controlling the emergence of zoonoses. Our immediate results surrounding the SARS-CoV-2 intermediate host are more suitable to providing context to an ongoing public health emergency, while the theory we develop provides a solid foundation for future research and reanalysis of previous and ongoing epidemics. Further, this model can provide insight into pathogens that are endemic in humans but still evolving in their intermediate host species, such as avian influenza. Our hope is that this theoretical framework can help predict the intermediate host species of future pathogens with specific ecological and epidemiological qualities, thus focusing research and intervention before a pathogen reaches human populations.

CRediT authorship contribution statement

Katherine Royce: Conceptualization, Methodology, Software, Validation, Formal analysis, Writing - original draft, Writing - review & editing.

Acknowledgments

The authors have no sources of funding to declare.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, athttps://doi.org/10.1016/j.jtbi.2021.110761.

References

Allen, L., Brown, V., Jonsson, C., Klein, S.L., Laverty, S., Magwedere, K., Owen, J., Van Den Driessche, P., 2012. Mathematical modeling of viral zoonoses in wildlife. Nat. Resour. Model. 25 (1), 5–51. https://doi.org/10.1111/j.1939-7445.2011.00104.x.

Banos, A., Lacasa, J. 2007 Spatio-temporal exploration of SARS epidemic, Cybergeo: European Journal of Geography doi:10.4000/cybergeo.12803..

Bolles, M., Donaldson, E., Baric, R., 2011. SARS-CoV and emergent coronaviruses: viral determinants of interspecies transmission. Curr. Opin. Virol. 1 (6), 624–634. https://doi.org/10.1016/j.coviro.2011.10.012.

Breban, R., Riou, J., Fontanet, A., 2013. Interhuman transmissibility of Middle East respiratory syndrome coronavirus: estimation of pandemic risk. Lancet 382 (9893), 694–699. https://doi.org/10.1016/S0140-6736(13)61492-0.

Chowell, G., Castillo-Chavez, C., Fenimore, P.W., Kribs-Zaleta, C.M., Arriola, L., Hyman, J.M., 2004. Model parameters and outbreak control for SARS. Emerg. Infect. Dis. 10 (7), 1258. https://doi.org/10.3201/eid1007.030647.

Corman, V.M., Muth, D., Niemeyer, D., Drosten, C. 2018. Hosts and sources of endemic human coronaviruses, in: Advances in Virus Research, Vol. 100, Elsevier, pp. 163–188. doi: 10.1016/bs.aivir.2018.01.001..

Cui, J., Li, F., Shi, Z.-L., 2019. Origin and evolution of pathogenic coronaviruses. Nat. Rev. Microbiol. 17 (3), 181–192. https://doi.org/10.1038/s41579-018-0118-9.

Cunningham, A.A., Daszak, P., Wood, J.L., 2017. One health, emerging infectious diseases and wildlife: two decades of progress?. Phil. Trans. Roy. Soc. B: Biol. Sci. 372 (1725), 20160167. https://doi.org/10.1098/rstb.2016.0167.

Dudas, G., Carvalho, L.M., Rambaut, A., Bedford, T., 2018. MERS-CoV spillover at the camel-human interface. eLife 7, e31257. https://doi.org/10.7554/eLife.31257.

Fan, Y., Zhao, K., Shi, Z.-L., Zhou, P., 2019. Bat coronaviruses in China. Viruses 11 (3), 210. https://doi.org/10.3390/v11030210.

Guan, Y., Zheng, B., He, Y., Liu, X., Zhuang, Z., Cheung, C., Luo, S., Li, P., Zhang, L., Guan, Y., et al., 2003. Isolation and characterization of viruses related to the SARS coronavirus from animals in southern China. Science 302 (5643), 276–278. https://doi.org/10.1126/science.1087139.

Hu, B., Zeng, L.-P., Yang, X.-L., Ge, X.-Y., Zhang, W., Li, B., Xie, J.-Z., Shen, X.-R., Zhang, Y.-Z., Wang, N., et al., 2017. Discovery of a rich gene pool of bat SARS-related coronaviruses provides new insights into the origin of SARS coronavirus. PLoS Pathog. 13, (11). https://doi.org/10.1371/journal.ppat.1006698 e1006698.

Ji, W., Wang, W., Zhao, X., Zai, J., Li, X., 2020. Cross-species transmission of the newly identified coronavirus 2019-nCoV. J. Med. Virol. 92 (4), 433–440. https://doi.org/10.1002/jmv.25682.

Lam, S.D., Bordin, N., Waman, V.P., Scholes, H.M., Ashford, P., Sen, N., Van Dorp, L., Rauer, C., Dawson, N.L., Pang, C.S., et al. 2020 SARS-CoV-2 spike protein predicted to form stable complexes with host receptor protein orthologues from mammals. Sci. Rep. 10:16471. doi: 10.1038/s41598-020-71936-5..

Latinne, A., Hu, B., Olival, K.J., Zhu, G., Zhang, L., Li, H., Chmura, A.A., Field, H.E., Zambrana-Torrelio, C., Epstein, J.H., et al., 2020. Origin and cross-species transmission of bat coronaviruses in china. Nat. Commun. 11 (1), 1–15. https://doi.org/10.1038/s41467-020-17687-3.

Lau, S.K., Woo, P.C., Li, K.S., Huang, Y., Tsoi, H.-W., Wong, B.H., Wong, S.S., Leung, S.-Y., Chan, K.-H., Yuen, K.-Y., 2005. Severe acute respiratory syndrome coronavirus-like virus in chinese horseshoe bats. Proc. Nat. Acad. Sci. 102 (39), 14040–14045. https://doi.org/10.1073/pnas.0506735102.

Li, W., Wong, S.-K., Li, F., Kuhn, J.H., Huang, I.-C., Choe, H., Farzan, M., 2006. Animal origins of the severe acute respiratory syndrome coronavirus: insight from ACE2-S-protein interactions. J. Virol. 80 (9), 4211–4219. https://doi.org/10.1128/JVI.80.9.4211-4219.2006.

Li, C., Yang, Y., Ren, L., 2019. Genetic evolution analysis of,novel coronavirus and coronavirus from other species. Infect. Genet. Evol. 2020,. https://doi.org/10.1016/j.meegid.2020.104285 104285.

Liu, P., Jiang, J.-Z., Wan, X.-F., Hua, Y., Li, L., Zhou, J., Wang, X., Hou, F., Chen, J., Zou, J., et al., 2020a. Are pangolins the intermediate host of the 2019 novel coronavirus (SARS-CoV-2)?. PLoS Pathog. 16, (5). https://doi.org/10.1371/journal.ppat.1008421 e1008421.

Liu, Z., Xiao, X., Wei, X., Li, J., Yang, J., Tan, H., Zhu, J., Zhang, Q., Wu, J., Liu, L., 2020b. Composition and divergence of coronavirus spike proteins and host ACE2 receptors predict potential intermediate hosts of SARS-CoV-2. J. Med. Virol. 92 (6), 595–601. https://doi.org/10.1002/jmv.25726.

Liu, J., Xie, W., Wang, Y., Xiong, Y., Chen, S., Han, J., Wu, Q., 2020c. A comparative overview of COVID-19, MERS and SARS. Int. J. Surg. 81, 1–8. https://doi.org/10.1016/j.ijsu.2020.07.032.

Lloyd-Smith, J.O., Funk, S., McLean, A.R., Riley, S., Wood, J.L., 2015. Nine challenges in modelling the emergence of novel pathogens. Epidemics 10, 35–39. https://doi.org/10.1016/j.epidem.2014.09.002.

Neumann, G., Noda, T., Kawaoka, Y., 2009. Emergence and pandemic potential of swine-origin H1N1 influenza virus. Nature 459 (7249), 931–939. https://doi.org/10.1038/nature08157.

Omrani, A.S., Al-Tawfiq, J.A., Memish, Z.A., 2015. Middle East respiratory syndrome coronavirus (MERS-CoV): animal to human interaction. Pathogens Global Health 109 (8), 354–362. https://doi.org/10.1080/20477724.2015.1122852.

Reusken, C.B., Haagmans, B.L., Müller, M.A., Gutierrez, C., Godeke, G.-J., Meyer,B., Muth, D., Raj, V.S., Smits-De Vries, L., Corman, V.M., et al. 2013 Middle East respiratory syndrome coronavirus neutralising serum antibodies in dromedary camels: a comparative serological study. Lancet Infectious Diseases 13 (10):859–866. doi:10.1016/S1473-3099(13)70164-6..

Royce, K., Fu, F., 2020. Mathematically modeling spillovers of an emerging infectious zoonosis with an intermediate host. PloS ONE 15, (8). https://doi.org/10.1371/journal.pone.0237780 e0237780.

Shi, Z., Hu, Z., 2008. A review of studies on animal reservoirs of the SARS coronavirus. Virus Res. 133 (1), 74–87. https://doi.org/10.1016/j.virusres.2007.03.012.

Shi, J., Wen, Z., Zhong, G., Yang, H., Wang, C., Huang, B., Liu, R., He, X., Shuai, L., Sun, Z., et al., 2020. Susceptibility of ferrets, cats, dogs, and other domesticated animals to SARS–coronavirus 2. Science 368 (6494), 1016–1020. https://doi.org/10.1126/science.abb7015.

Van den Driessche, P., Watmough, J., 2002. Reproduction numbers and sub-threshold endemic equilibria for compartmental models of disease transmission. Math. Biosci. 180 (1–2), 29–48. https://doi.org/10.1016/S0025-5564(02)00108-6.

Xu, J., Zhao, S., Teng, T., Abdalla, A.E., Zhu, W., Xie, L., Wang, Y., Guo, X., 2020. Systematic comparison of two animal-to-human transmitted human coronaviruses: SARS-CoV-2 and SARS-CoV. Viruses 12 (2), 244. https://doi.org/10.3390/v12020244.

Yuan, S., Jiang, S.-C., Li, Z.-L. 2020. Analysis of possible intermediate hosts of the new coronavirus SARS-CoV-2. Front. Veteri. Sci. 7. doi:10.3389/fvets.2020.00379..

Zhang, S., Diao, M., Yu, W., Pei, L., Lin, Z., Chen, D., 2020. Estimation of the reproductive number of novel coronavirus (COVID-19) and the probable outbreak size on the Diamond Princess cruise ship: A data-driven analysis. Int. J. Infect. Diseases 93, 201–204. https://doi.org/10.1016/j.ijid.2020.02.033.

Zhou, J., Li, C., Liu, X., Chiu, M.C., Zhao, X., Wang, D., Wei, Y., Lee, A., Zhang, A.J., Chu, H., et al., 2020. Infection of bat and human intestinal organoids by SARS-CoV-2. Nat. Med., 1–7 https://doi.org/10.1038/s41591-020-0912-6.